AD-A064 001   WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER   F/G 12/1

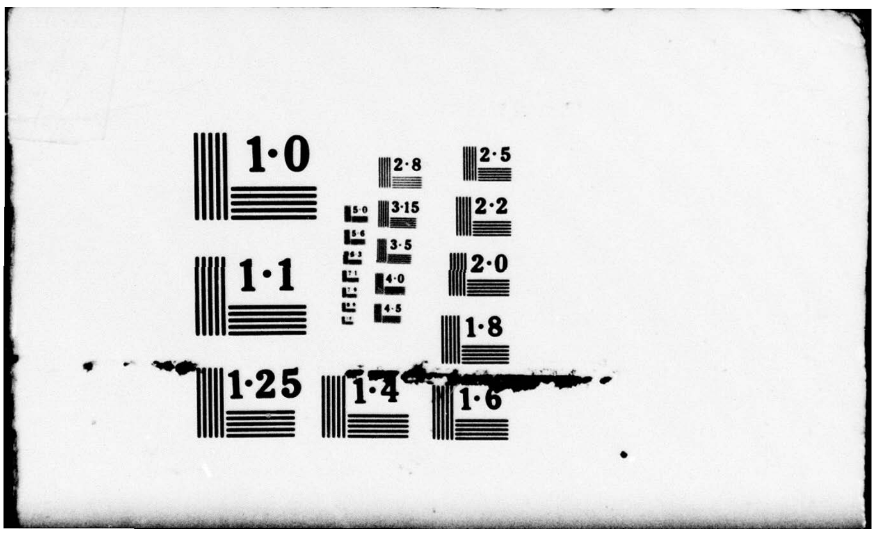
SHADOW PRICES, DUALITY AND GREEN'S FORMULA FOR A CLASS OF OPTIM--ETC(U)

NOV 78 J P AUBIN, F H CLARKE   DAAG29-75-C-0024

UNCLASSIFIED   MRC-TSR-1900   NL

1 OF 1
ADA
064001

END
DATE
FILMED
4 -79
DDC


1·0
2·8
2·5
5·0
3·15
2·2
5·6
3·5
1·1
2·0
4·0
4·5
1·8
1·25
1·4
1·6

MRC Technical Summary Report #1900

SHADOW PRICES, DUALITY AND GREEN'S FORMULA FOR A CLASS OF OPTIMAL CONTROL PROBLEMS

J. P. Aubin and F. H. Clarke

LEVEL

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

November 1978

Received September 20, 1978

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Research Council of Canada
Vancouver, B. C. Canada V6T 1W5

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

SHADOW PRICES, DUALITY AND GREEN'S FORMULA FOR A CLASS OF OPTIMAL CONTROL PROBLEMS

J. P. Aubin and F. H. Clarke

Technical Summary Report #1900
November 1978

ABSTRACT

A class of optimal control problems is considered in which the cost functional is locally Lipschitz (not necessarily convex or differentiable) and the dynamics linear and/or convex. By using generalized gradients and duality methods of functional analysis, necessary conditions are obtained in which the dual variables admit interpretation as shadow prices (or rates of change of the value function). Applications are presented in three settings: infinite-horizon optimal control, optimal control of partial differential equations, and a variational problem with unilateral state constraints. A theorem is proved which characterizes the generalized gradients of integral functionals on L^p.

L(P)

AMS(MOS) Subject Classification: 49B30

Key Words: maximum principle, sensitivity, infinite horizon, generalized gradients

Work Unit No. 5 - Mathematical Programming and Operations Research

---

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and the National Research Council of Canada under Grant No. A 9082.

page - i -

SIGNIFICANCE AND EXPLANATION

We develop in this paper a theory of optimality conditions for an abstract optimal control problem which encompasses a variety of situations that are met in practice, but not covered by the standard theory. Three such specific applications are made. The first of these deals with a control problem over an infinite interval, while the second and third involve the control of partial differential equations and state constraints, respectively. It is shown in each case that the general theory yields new and stronger necessary conditions, two significant features of which are the capability of treating nondifferentiable functions, and the interpretation of a certain quantity as the sensitivity of the problem with respect to certain perturbations.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

SHADOW PRICES, DUALITY AND GREEN'S FORMULA FOR A CLASS OF OPTIMAL CONTROL PROBLEMS

J. P. Aubin and F. H. Clarke

Introduction

We treat in this article a general class of optimal control problems in which the cost functional need only be locally Lipschitz, and in which the constraints exhibit linearity and convexity. We obtain for these hybrid problems necessary conditions with features usually associated only with the fully convex case. Chief among these is the interpretation of the multipliers as rates of change of the optimal value in the problem as the constraints undergo perturbation. In the context of mathematical programming such interpretations have figured importantly in mathematical economics (see for example [13] for a discussion), where the rates of change are in turn interpreted as "shadow prices". A further feature is the ability to guarantee the strong (Kuhn-Tucker or normal) form of the necessary conditions by means of a Slater-type constraint qualification independent of the solution to the problem. The method used here to treat these problems is novel, and employs the generalized gradient of the value function, a minimax theorem, and an abstract Green formula.

We present three applications of the abstract theory which we believe to be of independent interest; they are in settings characterized by technical difficulties. The first of these (§1) involves a control problem over an infinite interval, and sheds new light on the sensitive relationship between the growth rates of the cost functional and the adjoint variables, and on the role played by the size of the discount factor. In §2 we give an example to illustrate the importance of this consideration, while §3 gives the proof of the theorem in §1. We have placed this application before the abstract theory in order to display the line of reasoning common to both in a more concrete, and hence more easily assimilated setting.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and the National Research Council of Canada under Grant No. A 9082.

Section 4 recalls a stability result used in the proof, while §§5 and 6 develop characterizations of generalized gradients of certain functionals which are useful in interpreting the abstract conditions; these results complement other characterizations given in [9]. The abstract problem and its analysis appear in §§7 and 8. In §9 an application is made to a nondifferentiable, nonconvex problem involving control of partial differential equations. While this is the first such result that we know of, it should be clear that many other similar problems could be framed within the abstract theory. The final section illustrates the use of the theory in the presence of unilateral state constraints.

We conclude by recalling for the reader's convenience the definition of the generalized gradient in the case of a locally Lipschitz function $f: X \to R$, where X is a Banach space (details appear in [7] [9]). Given v, the generalized directional derivative $f^\circ(x;v)$ is defined by

$$f^\circ(x;v) = \limsup \, [f(y+\lambda v) - f(y)]/\lambda ,$$

where the upper limit is taken as y in X converges to x and λ in $(0,\infty)$ converges to 0. The generalized gradient of f at x, $\partial f(x)$, consists of all ζ in X^* such that

$$\langle v,\zeta\rangle \le f^\circ(x;v) \quad \text{for all } v \text{ in } X .$$

It follows that

$$f^\circ(x;v) = \max\{\langle v,\zeta\rangle : \zeta \in \partial f(x)\} ,$$

and that $\partial f(x)$ reduces to the derivative if f is continuously differentiable or to the subdifferential of convex analysis if f is convex.

1. An infinite-horizon optimal control problem.

We are given a locally Lipschitz function $g: R^n \times R^m \to R$, i.e., whenever (x,u), (y,v) are restricted to a suitably chosen neighborhood of any given point in $R^n \times R^m$, there is a constant K such that

$$|g(x,u) - g(y,v)| \le K|(x,u) - (y,v)| .$$

Also on hand are matrices F and G, $n \times n$ and $n \times m$ respectively, a point x_o in R^n, a positive number δ and a compact convex subset U of R^m.

The problem we consider is that of minimizing

$$(1.1) \qquad \int_0^\infty e^{-\delta t} g(x(t),u(t))dt$$

over the (so-called) trajectory-control pairs (x,u) which satisfy

$$(1.2) \qquad \begin{aligned} &u(t) \in U \text{ , a.e. } t \ge 0 \\ &\dot{x}(t) = Fx(t) + Gu(t) \text{ , a.e. } t \ge 0 \\ &x(0) = x_o \end{aligned}$$

where $u(\cdot)$ need only be measurable and x absol... y continuous.

We posit the following <u>growth condition</u> on g: there are numbers $c, r \ge 0$ such that, for every (x,u), for every ζ in the <u>generalized gradient</u> $\partial g(x,u)$ of g at (x,u), we have

$$(1.3) \qquad |\zeta| \le c(1+|(x,u)|^r) .$$

<u>Definition</u>. We denote by $\alpha(s)$ the infimum in the above problem when the initial condition is $x(0) = s$ (rather than $x(0) = x_o$), and by $\lambda(F)$ the maximum of the real parts of the eigenvalues of F.

We now suppose given a trajectory-control pair (x,u) which solves the above problem (with the original initial condition $x(0) = x_o$).

<u>Theorem 1</u>. Let (x,u) solve the problem as described above, and suppose $\delta > (r+1)\lambda(F)$. Then $\alpha(\cdot)$ is locally Lipschitz, and there exist an absolutely continuous function $p(\cdot)$ and a measurable function $(\zeta_1(\cdot),\zeta_2(\cdot))$ such that:

$$(1.4) \qquad -\dot{p}(t) = F^* p(t) - e^{-\delta t}\zeta_1(t) \text{ a.e.}$$

$$(1.5) \qquad (\zeta_1(t),\zeta_2(t)) \in \partial g(x(t),u(t)) \text{ a.e.}$$

$$(1.6) \qquad \max\{\langle p(t),Gw\rangle - e^{-\delta t}\langle w,\zeta_2(t)\rangle : w \in U\} = \langle p(t),Gu(t)\rangle - e^{-\delta t}\langle u(t),\zeta_2(t)\rangle \text{ a.e.}$$

$$(1.7) \qquad \int_0^\infty e^{(q-1)\delta t}|p(t)|^q dt < \infty , \int_0^\infty e^{(q-1)\delta t}|\dot{p}(t)|^q dt < \infty \text{ , where } q > 1$$

is defined by $1/q + 1/(r+1) = 1$, provided $r > 0$. If $r = 0$, we have instead: $e^{\delta t}|p(t)|$ and $e^{\delta t}|\dot{p}(t)|$ are bounded.

$$(1.8) \qquad \lim_{t\to\infty} e^{(q-1)\delta t}|p(t)|^q = 0 \text{ if } r > 0 \text{ ; if } r = 0 \text{ , then we have instead:}$$

$e^{\delta t}p(t)$ tends to a finite limit as t goes to $+\infty$.

$$(1.9) \qquad -p(0) \in \partial\alpha(x_o) .$$

<u>1.10 Remarks</u> (a) As shown in the proof, our hypotheses imply that the integral (1.1) is well-defined for all admissible pairs (x,u), so that no ambiguities result from our use of the word "solve".

(b) Except for the infinite interval of integration (called the case of an "infinite horizon" in economics) and the nondifferentiability of the cost functional, the above problem is a standard one in optimal control, and of the sort that arises often in mathematical economics. It is

in this connection especially that the interpretation of $p(0)$ as a marginal cost ("shadow price") associated with perturbing the initial condition, as afforded by (1.9), is useful. The roles of the infinite horizon and the discount factor are discussed in T.C. Koopmans [13]; see also [16].

(c) It is worth noting that in the above version of Pontryagin's maximum principle, the necessary conditions are stated in as strong a form as one could hope; i.e. "normally", without the presence of a possibly vanishing multiplier, and with strong "transversality conditions" at infinity. Similar statements have been derived (incorrectly in many instances) by reasoning by analogy with the finite horizon case. It was H. Halkin [11], to our knowledge, who first pointed out the perhaps unexpected pathology that can arise due to the infinite horizon. Based on the results of this paper, one could say that the necessary conditions may be expected to hold in the strong form provided the "discount rate" δ is sufficiently large. A further moral is that in the infinite-horizon case, growth (or dual) conditions such as (1.7) on the adjoint variable p are more natural than pointwise transversality conditions such as (1.8) (which are simply consequences of (1.7)). This fact was foreshadowed in Theorem A.8 of [4], which can be obtained as a corollary of Theorem 1. Finally, it is well-known (and true) that the conclusions of the theorem are sufficient for (x,u) to be optimal if g is jointly convex in (x,u).

2. An example

Consider the control system on R^2 given by

$$\dot{x}_1 = -x_2$$

$$\dot{x}_2 = -x_1 + u$$

$$-1 \le u \le 0$$

$$x_1(0) = 0 \ , \ x_2(0) = 0 \ ,$$

with the view of minimizing $\int_0^\infty e^{-\delta t} x_1(t)dt$. This is the case of the preceding section in which

$$F = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \ G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ U = [-1,0] \ .$$

It follows that x_1 satisfies

$$x_1(0) = \dot{x}_1(0) = 0 \ , \ \ddot{x}_1(t) \ge x_1(t) \ ,$$

so that, by a simple argument, $x_1(t) \ge 0$ for all $t \ge 0$. Consequently, the control $u \equiv 0$ and response $x \equiv 0$ are optimal for the problem.

The adjoint equation (1.4) becomes

$$-\dot{p}_1 = -p_2 - e^{-\delta t} \ , \ -\dot{p}_2 = -p_1 \ ,$$

and hence

$$\ddot{p}_2 = p_2 + e^{-\delta t} \ . \tag{2.1}$$

The condition (1.6) yields the fact that $\dot{x}_2(t)$ must equal -1 if $p_2(t)$ is negative. Thus we must have $p_2(t) \ge 0$ for all t.

The solution of (2.1) is of the form

$$p_2(t) = ae^t + be^{-t} + e^{-\delta t}/(\delta^2-1)$$

It is easy to see that if $0 < \delta < 1$ and $p_2(t)$ converges to 0 as $t \to \infty$ (as must be the case if (1.8) is to hold for any $r \geq 0$), then $p_2(t)$ is negative for large t. This contradiction is not inconsistent with the theorem, since in this case $\lambda(F) = 1$. Of course, if $\delta > 1$, the conclusions of the theorem hold.

Let us note finally that for $\delta > 1$, we may take $r = 0$ in applying the theorem, and it follows that $p_2(t)$ must be $e^{-\delta t}/(\delta^2-1)$. The condition (1.9) now gives

$$(\delta/(\delta^2-1),\ -1/(\delta^2-1)) \in \partial\alpha(0,0)\ ,$$

which makes evident the increasingly unstable dependence on the initial condition of the problem as δ nears 1.

3. <u>Proof of Theorem 1</u>. In this section, we denote the optimal pair $(\bar{x},\bar{u})$.

<u>STEP 1</u> We denote by μ the measure on $[0,\infty)$ having density $e^{-\delta t}dt$. We set $p = 1 + r$ and we let X be the space $W_n^{1,p} \times L_m^p$, where L_k^p means $L^p([0,\infty), R^k)$(with respect to μ) and where $W_n^{1,p}$ denotes the Sobolev space of absolutely continuous functions x such that x and $\dot{x}$ belong to L_n^p. (Recall that the norm of x in $W_n^{1,p}$ is given by $||x||_p + ||\dot{x}||_p$.) We set $Y = L_n^p \times L_m^p \times R^n$, and we define the subset Z of Y via

$$Z = \{0\} \times K \times \{x_o\}\ ,$$

where K is the set of functions u in L_m^p satisfying

$$u(t) \in U \quad \text{a.e.}$$

We define $f\colon X \to R$ by

$$f(x,u) = \int_0^\infty e^{-\delta t} g(x(t),\, u(t))dt\ ,$$

and $A\colon X \to Y$ by

$$A(x,u) = [\dot{x} - Fx - Gu,\ u,\ x(0)]\ .$$

Theorem 3 of §6 implies that f is locally Lipschitz on $L^p \times L^p$, and so all the more on $W_n^{1,p} \times L_m^p$. The optimal control problem of §1 may be phrased as that of minimizing $f(x,u)$ over X subject to $A(x,u) \in Z$. We shall denote the given solution to this problem $(\bar{x},\bar{u})$. Proposition 4.1 of §4 will be available to us as soon as (4.1) is verified, and the reader may check that this is equivalent to the following result: (we shorten $\lambda(F)$ to λ).

Lemma 3.1

If $\delta > p\lambda$ and v belongs to L_n^p, then the solution x to

$$\dot{x} = Fx + v\ ,\ x(0) = x_o$$

belongs to $W_n^{1,p}$. In fact, $||x|| \le \bar{c}||v||$ for some constant $\bar{c}$.

Proof It suffices to show that the function

$$x(t) = e^{Ft}x_o + \int_o^t e^{F(t-s)}v(s)ds$$

belongs to L_n^p. Since $|e^{Ft}|$ is bounded by $e^{\lambda t}$, it is clear that $e^{Ft}x_o$ belongs to L_n^p, so we need only study the last term. Now, choosing any q between λ and δ/p, and limiting ourselves to the case $r > 0$ (the case $r = 0$ calls for straightforward modifications), we have

$$\int_o^\infty e^{-\delta t}\left|\int_o^t e^{F(t-s)}v(s)ds\right|^p dt$$

$$\le \int_o^\infty e^{[\lambda p-\delta]t}\left[\int_o^t e^{s(q-\lambda)}e^{-sq}|v(s)|ds\right]^p dt$$

$$\le \int_o^\infty e^{[\lambda p-\delta]t}\left[\int_o^t e^{s(q-\lambda)p/r}ds\right]^r\left[\int_o^t e^{-sqp}|v(s)|^p ds\right]dt \qquad \text{(Hölder)}$$

$$\le \bar{c}\int_o^\infty e^{[\lambda p-\delta]t}\left[e^{t(q-\lambda)p}\right]\left[\int_o^t e^{-sqp}|v(s)|^p ds\right]dt$$

$$= \bar{c}\int_o^\infty e^{-sqp}|v(s)|^p ds\int_s^\infty e^{t(qp-\delta)}dt \qquad \text{(Fubini)}$$

$$= \bar{c}\int_o^\infty e^{-\delta s}|v(s)|^p ds < \infty\ .$$

Q.E.D.

Proposition 4.1 now assures us that the function α defined in §1 is Lipschitz in a neighborhood of x_o (in fact, it says rather more about a function involving more arguments). Note that if (x,u) in X is such that $\tilde{A}(x,u) \in \{0\} \times K$, where $\tilde{A}(x,u) = [\dot{x} - Fx - Gu,\ u]$, then

$$f(x,u) \ge \alpha(x(0))\ ,$$

with equality when $(x,u) = (\bar{x},\bar{u})$.

Thus $(\bar{x},\bar{u})$ minimizes

$$\tilde{f}(x,u) = f(x,u) - \alpha(x(0))$$

over X subject to $\tilde{A}(x,u) \in \{0\} \times K$. By redefining $\tilde{f}$ away from $(\bar{x},\bar{u})$, we can arrange to preserve this state of affairs while assuring that $\tilde{f}$ be locally Lipschitz on X. We may now apply Proposition 4.1 again to deduce that the function $\beta\colon L_n^p \times L_m^p \to R$ is Lipschitz in a neighborhood of $(0,0)$, where (B is the closed unit ball in R^n)

$$\beta(v,w) = \inf\{\tilde{f}(x,u)\colon A(x,u) \in \{0\} \times K + (v,w)\ ,\ x(0) \in x_o + B\}\ .$$

STEP 2 Let us now choose any (x,u) in X and w in K. Upon observing that for any $\epsilon > 0$ sufficiently small

$$\tilde{A}((\bar{x},\bar{u}) + \epsilon(x-\bar{x},u-\bar{u})) \in \{0\} \times K + \epsilon(\dot{\bar{x}}-F\bar{x}-Gu,u-w)\ ,$$

$$(\bar{x}+\epsilon(x-\bar{x}))(0) \in x_o + B\ ,$$

we deduce

$$\tilde{f}((\bar{x},\bar{u}) + \epsilon(x-\bar{x},u-\bar{u})) \ge \beta(\epsilon(\dot{x}-Fx-Gu,u-w))\ .$$

After subtracting $\tilde{f}((\bar{x},\bar{u})) = \beta(0,0)$ from either side, dividing by ϵ and

taking upper limits, we arrive at

$$(-\beta)^\circ(0,0;\, \dot{x}-Fx-Gu, u-w) + f^\circ(\bar{x},\bar{u}; x-\bar{x}, u-\bar{u})$$
$$+ (-\alpha)^\circ(x_o; x(0)-x_o) \geq 0 .$$

Recall [7][9] that $f^\circ(x;v)$ equals $\max\{\langle\zeta,v\rangle : \zeta \in \partial f(x)\}$. Consequently,

$$\min_{x,u,w}\ \max_{\gamma,\zeta,v}\ \{\langle\gamma_1,\, \dot{x}-Fx-Gu\rangle + \langle\gamma_2,\, u-w\rangle + \langle\zeta_1,\, x-\bar{x}\rangle$$
$$+ \langle\zeta_2,\, u-\bar{u}\rangle + \langle v,\, x(0)-x_o\rangle\} = 0 ,$$

where the max is over all (x,u,w) in $X \times K$ and the min over all $\gamma = (\gamma_1,\gamma_2)$, $\zeta = (\zeta_1,\zeta_2)$, and v lying in $\partial(-\beta)(0,0)$, $\partial f(\bar{x},\bar{u})$, and $\partial(-\alpha)(x_o)$ respectively. Because these sets are w^*-compact, the "lop-sided" minimax theorem [1] applies, and we deduce the existence of ζ,γ,v such that, for all (x,u) in X and w in K ,

$$\langle\gamma_1,\, \dot{x}-Fx-Gu\rangle + \langle\gamma_2, u-w\rangle + \langle\zeta_1,\, x-\bar{x}\rangle \tag{3.1}$$
$$+ \langle\zeta_2,\, u-\bar{u}\rangle + \langle v,\, x(0)-x_o\rangle \geq 0 .$$

<u>STEP 3</u> Let us set $w = u = \bar{u}$ in (3.1). We then deduce

$$\langle\gamma_1,\, \dot{x}-\dot{\bar{x}}-F(x-\bar{x})\rangle + \langle\zeta_1,\, x-\bar{x}\rangle + \langle v,\, x(0)-x_o\rangle = 0$$

for all x in $W_n^{1,p}$. Because $W_n^{1,p}$ is dense in L_n^p (this follows, for example, from the fact that $\mathcal{D}_n(0,\infty)$ is dense in L_n^p), it follows from Proposition 5.1, §5, that ζ_1 belongs to L_n^{p*} , the dual of L_n^p , rather than merely to the dual of $W_n^{1,p}$ (which is best avoided). Of course, ζ_2, γ_2 belong to L_m^{p*} , γ_1 to L_n^{p*} , and v to R^n . Thus we have, for all x in $W_n^{1,p}$,

$$\int_0^\infty e^{-\delta t}\gamma_1(t)\cdot(\dot{x}-Fx)dt + \int_0^\infty e^{-\delta t}\zeta_1(t)\cdot x(t)dt + v\cdot x(0) = 0 .$$

A classical, now familiar argument (Dubois-Reymond lemma) employing integration be parts derives from this the conclusion that γ_1 is absolutely continuous, and that

$$\frac{d}{dt}\{e^{-\delta t}\gamma_1(t)\} = -F^* e^{-\delta t}\gamma_1(t) + e^{-\delta t}\zeta_1(t) ,$$
$$\gamma_1(0) = v .$$

If we set $p(t) = e^{-\delta t}\gamma_1(t)$, then the above immediately yields (1.4), (1.7), (1.9), while (1.5) is a consequence of Theorem 2, §6 . Condition (1.8) is an elementary consequence of (1.7), the line of argument being that found in [4, Lemma A.7]. *It remains to prove* (1.6).

<u>STEP 4</u> Set $x = \bar{x}$, $w = \bar{u}$ in (3.1). Then

$$-\langle\gamma_1,\, G(u-\bar{u})\rangle + \langle\gamma_2,\, u-\bar{u}\rangle + \langle\zeta_2,\, u-\bar{u}\rangle = 0$$

for all u in L_m^p . This implies

$$e^{-\delta t}\gamma_2(t) = G^* p(t) - e^{-\delta t}\zeta_2(t) \quad \text{a.e.} \tag{3.2}$$

Now set $x = \bar{x}$, $u = \bar{u}$ in (3.1). We deduce, for every w in K ,

$$\langle\gamma_2,\, w-\bar{u}\rangle \leq 0 . \tag{3.3}$$

It follows that for almost each t , we have

$$\gamma_2(t)\cdot w \leq \gamma_2(t)\cdot \bar{u}(t) \quad \text{for all } w \text{ in } U$$

(for if this were false we could, from the measurable selection theorem,

find $w(\cdot)$ in K contradicting (3.3)). Recalling (3.2), we arrive at

$$(G^* p(t) - e^{-\delta t}\zeta_2(t)) \cdot (w-\bar{u}(t)) \le 0 ,$$

which is (1.6) . Q.E.D.

4. A stability result recalled

Let X and Y be Banach spaces and $A: X \to Y$ a continuous linear operator. Consider, for s in a neighborhood of 0 in Y, the family of optimization problems consisting of minimizing a given function $F: X \to R$ subject to the constraints

$$x \in \Omega$$
$$Ax \in Z + s ,$$

where Ω, Z are given subsets of X and Y respectively. We denote by $C(s)$ the set of x satisfying these constraints, and by $\alpha(s)$ the infimum in the above problem. Thus

$$\alpha(s) = \inf\{F(x): x \in C(s)\} .$$

The following, a special case of [3, Proposition 2], is stated here for the reader's convenience.

<u>Proposition 4.1</u> Let Z and Ω be closed and convex, and F locally Lipschitz. Suppose that there is a bounded subset K such that $C(s) \subset K$ when s is near 0, and such that F is Lipschitz on K. Then, if the condition

(4.1) $$0 \in \operatorname{int}(A\Omega - Z)$$

is satisfied, and if $\alpha(0) > -\infty$, the function α is Lipschitz near 0.

5. A result on generalized gradients

Let X and Y be Banach spaces such that X is continuously imbedded in Y and such that X, as a subset of Y, is dense. Let $f: Y \to R$ be a function which is locally Lipschitz. It follows that its restriction to X is locally Lipschitz in the norm of X. We denote by ∂f_X the generalized gradient (see below) (in X^*) of this restriction, and by ∂f_Y the generalized gradient (in Y^*) of the function f defined on Y.

<u>Proposition 5.1</u> Let x be a point of X. Then

$$\partial f_X(x) \subset \partial f_Y(x),$$

in the sense that every ζ in $\partial f_X(x)$ admits a unique extension to an element of $\partial f_Y(x)$.

<u>Proof</u>: Recall [7][9] the generalized directional derivative $f_Y^\circ(x;v)$, defined by

$$f_Y^\circ(x;v) = \lim\sup\{f(y+\lambda v) - f(y)\}/\lambda,$$

where the upper limit is taken as y converges to x in Y and λ decreases to 0 in R. The generalized gradient $\partial f_Y(x)$ consists by definition of those elements ζ in Y^* satisfying

$$f_Y^\circ(x;v) \geq \langle\zeta,v\rangle \quad \text{for all } v \text{ in } Y.$$

It follows easily from density and from the fact that convergence in X implies convergence in Y that

$$f_X^\circ(x;v) \leq f_Y^\circ(x;v)$$

whenever x and v lie in X. Now let ζ belong to $\partial f_X(x)$. Then (by

definition)

$$\langle\zeta,v\rangle \leq f_X^\circ(x,v) \quad \text{for all } v \text{ in } X,$$

and by the preceding, along with the fact that the function $v \to f_Y^\circ(x;v)$ is bounded on bounded subsets of Y, we deduce that the function

$$v \to \langle\zeta,v\rangle$$

mapping X, with its topology induced by Y, to R is bounded on bounded subsets. It follows that this linear function is continuous on X (with the induced topology) and hence, by a standard argument, admits a unique linear extension to the complete space Y in which X is dense. The extension, which we also label ζ, still satisfies

$$\langle\zeta,v\rangle \leq f_Y^\circ(x;v)$$

for all v in Y, and hence belongs to $\partial f_Y(x)$ by definition.

Q.E.D.

6. Generalized gradients on L^p

We are given a complete measure space $(T,\mathcal{T},\mu)$ with $\mu(T) < +\infty$, a separable Banach space X, and a function $g: T \times X \to R$. We assume that the mapping $t \to g(t,x)$ is measurable for each x, that $x \to g(t,x)$ is locally Lipschitz for each t, and we posit the existence of scalars $p \geq 1$ and $c \geq 0$ such that for every (t,x), for every element ζ of $\partial_x g(t,x)$ (generalized gradient in x for t fixed), the following bound holds:

$$|\zeta| \leq c(1 + |x|^{p-1}) . \tag{6.1}$$

Finally, we suppose that $t \to g(t,0)$ is (finitely) integrable, and we set, for x in $L^p(T,X)$,

$$F(x) = \int_T g(t,x(t))\mu(dt) . \tag{6.2}$$

As usual, p_* denotes the (possibly infinite) quantity satisfying

$$1/p + 1/p_* = 1 .$$

Theorem 2. Under the above hypotheses, the function $F: L^p \to R$ given by (6.2) is well-defined (finitely) and locally Lipschitz (in fact, Lipschitz on bounded subsets of L^p), and we have

$$\partial F(x) \subset \int_T \partial_x g(t,x(t))\mu(dt) .$$

This means that corresponding to any element ζ of $\partial F(x)$, there is a function $\zeta(\cdot)$ in $L^{p_*}(T,X^*)$ satisfying

$$\zeta(t) \in \partial_x g(t,x(t)) \qquad \mu\text{-a.e.}$$

and such that, for every y in $L^p(T,X)$,

$$\langle\zeta,y\rangle = \int_T \langle\zeta(t), y(t)\rangle\mu(dt) .$$

Proof. The growth condition (6.1) is easily seen to imply

$$|g(t,x)| \leq |g(t,0)| + c(|x| + |x|^p) ,$$

which, combined with the fact that $t \to g(t,x(t))$ is measurable, yields the fact that $F(x)$ is defined and finite whenever x lies in L^p. We now prove that F is locally Lipschitz. Invoking the mean value theorem for generalized gradients [14] we have

$$|F(x) - F(y)| = \left| \int_T \langle z(t), x(t) - y(t)\rangle\mu(dt\right.$$

where $z(t)$ belongs to $\partial_x g(t,w(t))$ and $w(t)$ lies between $x(t)$ and $y(t)$. Using (6.1) and Hölder's inequality we have this last expression bounded above by

$$c\int_T \{1 + (|x(t)|+|y(t)|)^{p-1}\}|x(t) - y(t)|\mu(dt)$$

$$\leq c_1||x-y||_p + c||(|x|+|y|)^{p-1}||_{p_*}||x-y||_p$$

$$\leq (c_1 + c(||x||_p+||y||_p)^{p/p_*} ||x-y||_p$$

$$\leq K||x-y||_p \text{ , for some constant } K ,$$

as long as x and y remain in a bounded subset of L^p.

Now let ζ belong to $\partial F(x)$ (and hence to the dual of $L^p(T,X)$, $L^{p_*}(T,X^*)$ [5, §2.6, Prop. 10 and No. 21]). Then for any v in L^p,

$F^\circ(x;v) \geq \langle \zeta, v \rangle$. Using Fatou's Lemma (cf. [8, Lemma 3]) we may show

$$\int_T g_x^\circ(t,x(t);\, v(t))\mu(dt) \geq F^\circ(x;v) \ ,$$

from which ensues

$$\int_T g_x^\circ(t,x(t);\, v(t))\mu(dt) \geq \langle \zeta, v \rangle = \int_T \langle \zeta(t),\, v(t) \rangle \mu(dt) \ . \tag{6.3}$$

For any $\varepsilon > 0$, define a multifunction Γ as follows:

$$\Gamma(t) = \{0\} \quad \text{if} \quad g_x^\circ(t,x(t);v) > \langle \zeta(t), v \rangle - \varepsilon \quad \text{for all} \quad v \quad \text{in} \quad X \ ,$$
$$= \{v:\ g_x^\circ(t,x(t);v) \leq \langle \zeta(t), v \rangle - \varepsilon\} \quad \text{otherwise} \ .$$

The map $(t,v) \to g_x^\circ(t,x(t);v)$ is of course continuous in v , and is seen to be measurable in t as a consequence of the fact that $g_x^\circ(t,x(t);v)$ can be expressed as the upper limit of a countable family of measurable functions (we use here the fact that X is separable).

It follows that the multifunction Γ is "measurable" and admits a measurable selection $v(t)$ [19, Theorem 4.1]. Now (6.3) implies that the set

$$\{t:\ \Gamma(t) \neq \{0\}\}$$

must have μ-measure 0 , and since ε is arbitrary we deduce that for μ-almost all t , for all v in X ,

$$g_x^\circ(t,x(t);v) \geq \langle \zeta(t), v \rangle \ .$$

Consequently $\zeta(t)$ belongs to $\partial_x g(t,x(t))$ μ-a.e.

Q.E.D.

<u>Remark.</u> The case $p = \infty$ is treated in [9]. As in that case, equality will hold in the theorem if $g(t,\cdot)$ is "regular" for μ-almost all t .

7. An abstract approach

A. The line of reasoning used to prove Theorem 1 will work in much more general situations, as we now show by considering the following abstract optimization problem. We are given Banach spaces U,V,W,T,Z together with continuous linear operators $L: W \to V$, $\gamma: W \to T$, $C: U \to Z$, locally Lipschitz functions $f: U \to R$ and $g: T \to R$, and a multifunction E mapping U to V (i.e. $E(x)$ is a subset of V for x in U). W, the domain of L, is assumed to be a subset of U.

We consider the problem of minimizing

$$f(x) + g(\gamma x) \tag{7.1}$$

over the elements x of W which satisfy

$$Lx \in E(x) \tag{7.2}$$

$$\gamma x \in S \tag{7.3}$$

$$Cx \in Y, \tag{7.4}$$

where S and Y are specified subsets of T and Z respectively. Here, L plays the role of a differential operator and γ that of a trace operator, so that (7.1) is a type of Bolza functional, (7.2) is a "differential inclusion", (7.3) a boundary condition, and (7.4) an explicit "state constraint". An implicit constraint is also incorporated by (7.2), which demands that x belong to the domain of E; i.e. the set of points x for which $E(x) \neq \emptyset$.

We posit the following conditions:

(H_1) The sets S, Y, and $Gr(E)$ are closed and convex ($Gr(E)$, the graph of E, is the set of points (x,v) such that v belongs to $E(x)$.)

(H_2) ("trace property") The injection from W into U is continuous, γ has a continuous right inverse, and the kernel W_o of γ is dense in U.

Remark 7.5 The type of setting we have described above is most familiar in partial differential equations. The point as far as we are concerned is, first, to avoid the dual W^* of W (typically nasty) and have intervene instead the dual U^* of U, and, second, to avoid the use of the transpose L^* of L and to replace it by the transpose $L_o^* \in L(V^*, W_o^*)$ of the restriction $L_o \in L(W_o, V)$ of L to W_o. The motivation stems from the familiar case in which L is a differential operator and W is a space of functions or distributions; when W_o is also the closure of the space of infinitely differentiable functions with compact support, then W_o^* is a subspace of distributions, and L_o^* is also a differential operator (in the sense of distributions) and can be computed, whereas the transpose L^* cannot in general be expressed in terms of differential operators.

When the trace property (H_2) holds, we can compare L^* and L_o^* by means of an abstract Green formula in the following way (see Aubin [2]). First we introduce the domain V_o^* of L_o^* defined by

$$V_o^* = \{p \in V^* : L_o^* p \in U^*\}, \tag{7.6}$$

where U^* is identified with a subspace of W_o^* (indeed, the transpose of the injection from W_o into U is an injection from U^* to W_o^*). Equipped with the graph norm, V_o^* is a Banach space, and if (H_2) holds then there exists a unique operator $\beta^* \in L(V_o^*, T^*)$ such that

$$\left\{ \begin{array}{l} \text{for all } x \text{ in } W, \text{ for all } p \text{ in } V_o^*, \\ \langle L_o^* p, x\rangle - \langle p, Lx\rangle = \langle \beta^* p, \gamma x\rangle . \end{array} \right. \tag{7.7}$$

(This will replace the integration by parts in the proof of Theorem 1.)

We associate to all (u,v,t,z) in $U \times V \times T \times Z$ the set

$\Gamma(u,v,t,z)$ of points x in W satisfying

(7.8) $$Lx \in E(x-u) + v\ ,\ \gamma x \in S + t\ ,\ Cx \in Y + z\ .$$

Notice that our optimization problem involves minimizing a functional over $\Gamma(0,0,0,0)$. We make the following controllability assumption:

(H_3) There is a bounded subset Γ_o of U such that, for every (u,v,t,z) sufficiently small, the set $\Gamma(u,v,t,z)$ is nonempty and contained in Γ_o. The functions f and g are Lipschitz on neighborhoods of Γ_o and $\gamma(\Gamma_o)$ respectively.

We define the <u>value function</u> α on $U \times V \times T \times Z$ via

$$\alpha(u,v,t,z) = \inf\{f(x) + g(\gamma x) : x \in \Gamma(u,v,t,z)\}\ ,$$

and the <u>Hamiltonian</u> function H on $U \times V^*$ via

$$H(x,p) = \sup\{\langle p,v\rangle : v \in E(x)\}\ .$$

Since the graph of E is convex, H is concave with respect to x and convex with respect to p. We denote by $\partial_p H(x,p)$ the subdifferential in the sense of convex analysis [17] of the convex function $p \to H(x,p)$, and by $\partial_x H(x,p)$ the superdifferential of the concave function $x \to H(x,p)$.

<u>Theorem 3</u> We posit (H_1)-(H_3). If x in W minimizes (7.1) subject to the constraints (7.2)-(7.4), there exist $p \in V_o^*$, $\psi \in \partial f(x)$, $\phi \in \partial g(\gamma x)$, and $\tau \in N_Y(C\ )$ satisfying

(7.9) $$L_o^* p \in \partial_x H(x,p) - \psi - C^*\tau \subset \partial_x H(x,p) - \partial f(x) - C^* N_Y(Cx)$$

(7.10) $$Lx \in \partial_p H(x,p)$$

(7.11) $$\beta^* p \in \phi + N_S(\gamma x) \subset \partial g(\gamma x) + N_S(\gamma x)\ .$$

Futhermore, the value function α is Lipschitz in a neighborhood of $(0,0,0,0)$, and we have:

(7.12) $$(L_o^* p + C^*\tau + \psi,\ -p,\ -\beta^* p + \phi,\ -\tau) \in \partial\alpha(0,0,0,0)\ .$$

<u>Remark</u>. We recognize (7.9)-(7.10) to be Hamiltonian equations and (7.11) a transversality condition. The notation $N_S(\gamma x)$, for example, refers to the normal cone to the convex set S at the point γx(i.e. the set of p in T^* such that $\langle p, s-\gamma x\rangle \le 0$ for all s in S).

B. The shadow price information (7.12) exists with respect to four perturbations in the problem. In some situations it may be deemed more natural or more convenient to consider instead a performance function α depending on fewer variables. For example, in the problem of §1, we chose to mention only the interpretation of $p(0)$. One can easily imagine situations (e.g. sensitivity analysis) in which the generality of (7.12) would come into play in other ways.

As far as reduced versions of (7.12) is concerned, one cannot simply drop components to obtain them, since it is not generally true that the relation $(a,b) \in \partial f(x,y)$ implies $a \in \partial_x f(x,y)$ (this is discussed in [9, art. 14]). However, one can modify the proof of Theorem 3 to attain the desired result. We shall not attempt to consider the most general situation, but rather one that seems natural in many settings. We consider the case in which S is a singleton set $\{s_o\}$ and g is identically zero (i.e. boundary conditions reduce to $\gamma x = s_o$). We let $\delta(s)$ be the infimum in the problem in which the boundary condition is $\gamma x = s$, and we maintain the hypotheses of the theorem.

Corollary 4. If x is optimal for the above problem, there exist p, ψ and τ as in Theorem 3 such that (7.9)(7.10) hold. Further, $\mathfrak{d}$ is Lipschitz near s_o, and

$$-\beta^* p \in \partial \mathfrak{d}(s_o) \,. \tag{7.13}$$

C. We consider now a different version of the above problem, in which there is an explicit dependence on a control parameter, and unilateral constraints on the state variable. We introduce a Banach space Σ, closed convex subsets K of Σ and Ω of W, and linear operators F and G in $L(U,V)$ and $L(\Sigma,V)$ respectively. We now seek to minimize

$$f(x,\sigma) + g(\gamma x) \tag{7.14}$$

subject to

$$\begin{aligned} x &\in \Omega \\ Lx &= Fx + G\sigma \\ \sigma &\in K \\ \gamma x &\in S \\ Cx &\in Y \,, \end{aligned}$$

where $f: U \times \Sigma \to R$, and g, γ, L, S, C, Y are unchanged from Theorem 3. We denote by $\Gamma(u,s,v,t,z)$ the set of (x,σ) in $W \times \Sigma$ satisfying

$$\begin{aligned} x &\in \Omega + u \\ Lx &= Fx + G\sigma + v \\ \sigma &\in K + s \\ \gamma x &\in S + t \\ Cx &\in Y + z \,, \end{aligned}$$

and we define $\mathfrak{a}$ on $U \times \Sigma \times V \times T \times X$ via

$$a(u,s,v,t,z) = \inf\{f(x,\sigma) + g(\gamma x) : (x,\sigma) \in \Gamma(u,s,v,t,z)\} \,.$$

(H_4) There is a bounded subset Γ_o of $W \times \Sigma$ such that, for every (u,s,v,t,z) sufficiently small, $\Gamma(u,s,v,t,z)$ is non-empty and contained in Γ_o. Furthermore, f is Lipschitz on a neighborhood of Γ_o.

Corollary 5. Under the stated assumptions, if (x,σ) solves the above problem, there exist $p \in V_o^*$, (ψ_1,ψ_2) in $\partial f(x,\sigma)$, $\phi \in \partial g(\gamma x)$, and $\tau \in N_Y(Cx)$ satisfying

$$-L_o^* p + F^* p \in \psi_1 + C^* \tau + N_\Omega(x) \tag{7.15}$$

$$G^* p \in \psi_2 + N_K(\sigma) \tag{7.16}$$

$$\beta^* p \in \phi + N_S(\gamma x) \tag{7.17}$$

$$(L_o^* p + C^* \tau + \psi_1, \psi_2, -p, -\beta^* p + \phi, -\tau) \in \partial a(0,0,0,0,0) \,. \tag{7.18}$$

Remark 7.19: As before, it is possible to replace (7.18) be alternate relations, such as (7.13) when the boundary conditions are simply $\gamma x = s_o$. The reader may wish to verify that Theorem 1 is a consequence of this Corollary with the identifications (see §3 for notation):

$$W = \Omega = W_n^{1,p} \,,\; \Sigma = L_m^p \,,\; U = V = L_n^p$$

$$K = \{\sigma \in \Sigma : \sigma(t) \in U \text{ a.e.}\}$$

$$f(x,\sigma) = \int_0^\infty e^{-\delta t} g(x(t),\sigma(t))\,dt$$

$$Lx = \dot{x} \,,\; \gamma x = x(0)$$

$$Z = W = Y, C = \text{identity}$$

$$L_o^* p = -\dot{p}, \ \beta^* p = p(0)$$

$$V_o^* = \{p \in L_n^q : \dot{p} \in L_n^q\} .$$

Of course, in applying the theorem to specific problems, there will be a need for appropriate characterizations of generalized gradients (such as Theorem 2 of §6) and normal cones (such as the lemma in §10). A further element will be the verification of the controllability hypothesis (H_3). In the problem of §1, this was a result of taking δ large enough; in the applications of §§9 and 10, it follows from postulating "strictly feasible points", a concept akin to the Slater condition of mathematical programming.

8. Proof of the abstract necessary conditions.

We shall merely sketch the proof, since the steps are identical to those in the proof of Theorem 1, albeit in a more general setting. We define, in the notation of §4,

$$F(x) = f(x) + g(\gamma x)$$

$$X = W$$

$$Y = U \times V \times T \times Z$$

$$Z = Gr(E) \times S \times Y$$

$$Ax = (x, Lx, \gamma x, Cx) .$$

Then (H_3) and Proposition 4.1 imply that the function α is Lipschitz in a neighborhood of 0. As in §3, we denote the optimal solution $\bar{x}$, and we note that for any θ in $(0,1)$, for any x in W and (u,v,t,z) in $Gr(E) \times S \times Y$, we have

$$\bar{x} + \theta(x-\bar{x}) \in \Gamma(\theta[x-u, Lx-v, \gamma x-t, Cx-z]) .$$

Consequently,

$$\alpha(\theta[u,v,t,z]) \le F(\bar{x}+\theta(x-\bar{x})) ,$$

and taking generalized directional derivatives leads to

$$0 \le F^\circ(\bar{x}; x-\bar{x}) + \alpha^\circ(0, u-x, v-Lx, t-\gamma x, z-Cx) .$$

Applying just as in §3 the lop-sided minimax theorem, we deduce the existence of $(q,-p,r,-\tau) \in \partial\alpha(0,0,0,0)$, $\psi \in \partial f(\bar{x})$, $\phi \in \partial g(\gamma\bar{x})$ such that, for all x in W and (u,v,t,z) in $Gr(E) \times S \times Y$,

$$(8.1) \qquad 0 \le \langle q, u-x\rangle - \langle p, v-Lx\rangle + \langle r, t-\gamma x\rangle - \langle \tau, z-Cx\rangle$$
$$+ \langle \phi, x-\bar{x}\rangle + \langle \phi, \gamma x-\gamma\bar{x}\rangle .$$

If we set $(u,v,t,z) = (\bar{x}, L\bar{x}, \gamma\bar{x}, C\bar{x})$ in (8.1), we obtain that for every x in W,

$$(8.2) \qquad \langle q-\phi, \bar{x}-x\rangle - \langle p, L(\bar{x}-x)\rangle + \langle r-\phi, \gamma(\bar{x}-x)\rangle - \langle \tau, C(\bar{x}-x)\rangle = 0 .$$

As x ranges over $\bar{x} + W_o$, this implies

$$q = L_o^* p + C^* \tau + \psi .$$

Now by definition, $(q,-p,r,-\tau)$ belongs to $\partial\alpha(0,0,0,0)$ and hence to $U^* \times V^* \times T^* \times Z^*$, while by the result of §5, ψ belongs to U^* (rather than merely to W^*); of course $C^*\tau$ belongs to U^*. These facts and the preceding equation imply that $L_o^* p$ belongs to U^*; i.e., that p belongs to V_o^*. Consequently we can use the Green formula (7.7) in (8.2) to deduce that for any x in W,

$$\langle r-\phi+\beta^* p, \gamma x\rangle = 0 .$$

Since γ is surjective, this yields

$$r = \phi - \beta^* p ,$$

and now (7.12) follows.

If in (8.1) we take $x = \bar{x}$, it follows immediately that

$$(-q,p) \in N_{Gr(E)}(\bar{x}, L\bar{x}), \ -r \in N_S(\gamma\bar{x}), \ \tau \in N_Y(C\bar{x}) .$$

It follows from convex analysis that $(-q,p)$ lies in $N_{Gr(E)}(\bar{x}, L\bar{x})$ iff

$$q \in \partial_x H(\bar{x},p), \ L\bar{x} \in \partial_p H(\bar{x},p) ,$$

and all the conclusions of the theorem ensue.

Q.E.D.

The proof of Corollary 4 uses the device introduced in §3: we observe that $\bar{x}$ minimizes

$$f(x) - \delta(\gamma x)$$

subject to

$$Lx \in E(x), \ Cx \in Y, \ \gamma x \in s_o + \delta B ,$$

where $\delta > 0$ is small, and we apply Theorem 3 to this new problem. The proof of Corollary 5 consists of applying Theorem 3 after the following relabelings:

$$U = U \times \Sigma, \ W = W \times \Sigma, \ V = V$$

$$f(x,\sigma) = f(x), \ L(x,\sigma) = Lx, \ C(x,\sigma) = Cx, \ \gamma(x,\sigma) = \gamma x$$

$$E(x,\sigma) = Fx + G\sigma \quad \text{if} \quad x \in \Omega \quad \text{and} \quad \sigma \in K$$
$$= \phi \qquad \text{otherwise.}$$

9. An example in partial differential equations.

Let Ω be a bounded open subset of R^n whose boundary is a smooth differential manifold. We introduce functionals f and g defined by

$$f(x,\sigma) = \int_\Omega \phi(w,x(w),\ \sigma(w))dw$$

$$g(\xi) = \int_\Gamma \theta(w,\xi(w))dm(w)$$

where $dm(w)$ is a measure on Γ.

We consider the solutions $x \in H^2(\Omega)$ to the Dirichlet problem for the Laplacian:

(9.1)
 i) $-\Delta x + x = \sigma$ (σ ranges over $L^2(\Omega)$)
 ii) $x|_\Gamma = 0$

We denote by $\frac{\partial}{\partial n}$ the normal derivative. Let Λ be a compact set and let c be a function $(w,\lambda) \in \Omega \times \Lambda \to c(w,\lambda) \in R$ satisfying

(9.2)
 i) $\lambda \in \Lambda$, $w \to c(w,\lambda)$ belongs to $L^2(\Omega)$
 ii) for almost all $w \in \Omega$, $\lambda \to c(w,\lambda)$ is continuous.

We consider the following problem:

Minimize $\int_\Omega \phi(w,x(w),\ \sigma(w))dw + \int_\Gamma \theta(w,\ \frac{\partial x}{\partial n}(w))dm(w)$

subject to the constraints (9.1) and

(9.3) σ belongs to the unit ball of $L^2(\Omega)$

(9.4) $\forall\ \lambda \in \Lambda\ ,\ \int_\Omega c(w,\lambda)x(w)dw \geq b(\lambda)$

where $\lambda \to b(\lambda)$ is a continuous function. We posit the following controllability assumption.

(9.5) There exist $\sigma_o \in L^2(\Omega)$ such that $||\sigma_o||_{L^2(\Omega)} \leq 1$ and such that the solution x_o of the Dirichlet problem $-\Delta x_o + x_o = \sigma_o$ and $x_o|_\Gamma = 0$ satisfies:

$$\forall\ \lambda \in \Lambda,\ \int_\Omega c(w,\lambda)x_o(w)dw > b(\lambda)\ .$$

Finally, we assume that the functions ϕ and θ satisfy

(9.6)
 i) the functions $w \to \phi(w,x,\sigma)$ and $\theta(w,\xi)$ are measurable for each x, σ and ξ
 ii) the functions $(x,\sigma) \to \phi(w,x,\sigma)$ and $\xi \to \theta(w,\xi)$ are locally Lipschitz for almost all w.
 iii) there exists a constant $c > 0$ such that
 $\partial\phi(w,x,\sigma) \subset c(1+|x|+|\sigma|)B$
 $\partial\theta(w,\xi) \subset c(1+|\xi|)B$
 iv) the functions $w \to \phi(w,0,0)$ and $\theta(w,0)$ are (finitely) integrable.

Theorem 5. We posit assumptions (9.2), (9.5) and (9.6). Let $(\bar{x},\bar{\sigma})$ be an optimal solution. Then there exist $p \in L^2(\Omega)$ whose Laplacian Δp belongs to $L^2(\Omega)$ and functions $\psi_1,\psi_2 \in L^2(\Omega)$ satisfying

$$(\psi_1(w),\psi_2(w)) \in \partial\phi(w,\bar{x}(w),\bar{\sigma}(w)) \quad \text{for almost all } w \in \Omega$$

and a non-negative Radon measure $\bar{\mu}$ on Λ satisfying

$$\int_\Lambda \left[\int_\Omega (c(w,\lambda)\bar{x}(w))dw-b(\lambda)\right]d\bar{\mu}(\lambda) = 0$$

such that

i) $-\Delta p + p + \psi_1 = \int_\Lambda c(\cdot,\lambda)d\bar{\mu}(\lambda)$

ii) $-p|_\Gamma \in \partial\theta(w, \frac{\partial\bar{x}(w)}{\partial n})$

and

$$\int_\Omega \langle p(w) - \psi_2(w), \bar{\sigma}(w)\rangle = ||p-\psi_2||_{L^2(\Omega)} .$$

<u>Remark</u>. Among the possible shadow price interpretations we could add to the above is the following: if $\alpha(\xi)$ is the infimum in the above problem when the boundary condition is $x|_\Gamma = \xi$ (instead of $x|_\Gamma = 0$), then α is Lipschitz on a neighborhood of 0 in $H^{3/2}(\Gamma)$ and:

$$-\frac{\partial p}{\partial n} \in \partial\alpha(0)$$

<u>Proof</u>. We use Corollary 5 in the following case

$U = V = \Sigma = L^2(\Omega)$, $W = H^2(\Omega)$, $\Omega = U$,

$T = H^{3/2}(\Gamma) \times H^{1/2}(\Gamma)$, $F = 0$, G = identity, K = unit ball in Σ,

γ is the operator defined by $\gamma x(w) = (x(w)|_\Gamma , \frac{\partial x(w)}{\partial n})$.

The trace theorem [15] implies that γ is surjective and that $\mathrm{Ker}\,\gamma = H_o^2(\Gamma)$, the closure of $\mathcal{D}(\Omega)$ in $H^2(\Omega)$. We define L by $Lx = -\Delta x + x$. Hence L_o^* is defined by $L_o^* p = -\Delta p + p$ (in the sense of distributions) and its domain is the space $H^0(\Omega,\Delta)$ of functions $p \in L^2(\Omega)$ such that $\Delta p \in L^2(\Omega)$.

The Green formula can be written

$$(9.7)\quad \begin{cases} \langle -\Delta p + p, x\rangle - \langle p, -\Delta x + x\rangle \\ = \int_\Gamma \frac{\partial}{\partial n} p(w)x(w)dm(w) - \int_\Gamma p(w) \frac{\partial}{\partial n} x(w)dm(w) \end{cases}$$

for smooth functions. By the abstract Green formula it still holds true when $x \in H^2(\Omega)$, $p \in H^o(\Omega,\Delta)$ since the operator β^* defined by

$$\beta^* p(w) = (\frac{\partial}{\partial n} p(w), -p(w))$$

can be extended in a unique way to a continuous linear operator β^* from $H^o(\Omega,\Delta)$ to $H^{-3/2}(\Gamma) \times H^{-1/2}(\Gamma)$ is such a way that formula (9.7) holds. Finally, we choose $S = \{0\} \times H^{1/2}(\Gamma)$, and the (isoperimetric) constraints are defined by the Banach space $Z = C(\Lambda)$, the subset $Y = b + C_+(\Lambda)$ and the map C defined by $Cx(\lambda) = \int_\Omega c(w,\lambda)\, x(w)\, dw$.

We now show that the controllability assumption (H_4) of Corollary 5 is satisfied. Indeed, the operator $x \to (-\Delta x+x, x|_\Gamma)$ is an <u>isomorphism</u> from $H^2(\Omega)$ onto $L^2(\Omega) \times H^{3/2}(\Gamma)$. Let M_1 be the norm of its inverse, $M_2 = \sup_{\lambda\in\Lambda} ||C(\cdot,\lambda)||_{L^2(\Omega)}$, $M_3 = \min_{\lambda\in\Lambda} [\int_\Omega c(w,\lambda)\, x_o(w)dw - b(\lambda)] > 0$. We choose $\gamma > 0$ such that $\gamma < M_3/(1+3M_1M_2)$. Now, if $||v||_{L^2(\Omega)}$, $||s||_{L^2(\Omega)}$, $||z||_{C(\Lambda)}$, $||\xi||_{H^{3/2}(\Gamma)}$; are less than γ , then the solutions $(x,\sigma) \in H^2(\Omega)$ of

i) $-\Delta x + x = \sigma + v$

ii) $x|_\Gamma = \xi$

iii) $\sigma \in K + s$

iv) $Cx \geq b + z$

satisfy $||\sigma|| < 1 + \gamma$ and $||x|| < M_1(1+3\gamma)$. Further, such solutions exist; for take $\sigma = \sigma_o + s$. Then the solution x to $-\Delta x + x = \sigma + v$ satisfies $||x - x_o|| < 3\gamma M_1$, and hence

$$Cx = Cx_o - b + C(x-x_o) + b$$
$$\geq M_3 - M_2||x - x_o|| + b \geq M_3 - M_2 3\gamma M_1 + b$$
$$> \gamma + b > b + z .$$

This yields (H_4) in our present setting.

Now, assumptions (9.6) on the functions ϕ and θ imply that the functionals f and g are Lipschitz on bounded subsets of $L^2(\Omega)$ and $L^2(\Gamma)$ respectively (see §6).

The assumptions of Corollary 5 are therefore satisfied. It remains to interpret the conclusions. There exist $p \in H^*(\Omega,\Delta)$, ψ_1 and $\psi_2 \in L^2(\Omega)$ such that, by theorem 2,

$$(\psi_1(w), \psi_2(w)) \in \partial\phi(w, \bar{x}(w), \bar{\sigma}(w)) \quad \text{for almost all } w \in \Omega$$

$\xi \in L^2(\Gamma)$ such that

$$\xi(w) \in \partial\theta(w, \frac{\partial\bar{x}}{\partial n}(w)) \quad \text{for almost all } w \in \Gamma \text{ and a nonnegative}$$

Radon measure $\bar{\mu} \in C(\Lambda)^*$ (i.e. $-\tau$ of Corollary 5) such that

$$\int_\Lambda \left[\int_\Omega c(w,\lambda)\bar{x}(w)dw - b(\lambda) \right] d\bar{\mu}(\lambda) = 0$$

They satisfy the following equations:

i) $-\Delta p + p + \psi_1 = \int_\Lambda c(\cdot,\lambda)d\bar{\mu}(\lambda)$

ii) $-p|_\Gamma = \xi$

iii) $\int_\Omega \langle p(w) - \psi_2(w), \bar{\sigma}(w)\rangle = ||p-\psi_2||_{L^2(\Omega)}$

Q.E.D.

10. Example - a variational problem with unilateral constraints.

As a final example, we consider the following problem in the calculus of variations: to minimize

$$\int_0^1 g(x(t), \dot{x}(t))dt$$

over the absolutely continuous arcs $x: [0,1] \to R^n$ satisfying

(10.1) $\quad x(t) \in \Omega$

(10.2) $\quad \dot{x}(t) \in K$ a.e.

(10.3) $\quad x(0) = x_o$, $x(1) = x_1$,

where Ω and K are given closed convex subsets of R^n , x_o and x_1 are given points in R^n , and $g: R^n \times R^n \to R$ is a locally Lipschitz function. We make the following assumptions: K is compact,
and there is an arc $x_o(\cdot)$ joining x_o to x_1 and an $\epsilon > 0$ such that $\dot{x}_o(t) + \epsilon B \subset K$ a.e. and such that $x_o(t)$ belongs to the <u>interior</u> of Ω for all t . We denote by $\alpha(t)$ the infimum in the above problem when the boundary conditions, instead of (10.3), are given by $x(0) = x_o$, $x(1) = t$.

<u>Theorem 6</u> If $\bar{x}$ solves the above problem, then α is Lipschitz near x_1 and there exist an absolutely continuous function p , an element λ of L^∞ , and a nonnegative Radon measure m on $[0,1]$ such that ($\int_t^1$ denotes integration over $[t,1]$):

(10.4) $\quad \lambda(t) \in N_\Omega(\bar{x}(t))$ m-a.e.

(10.5) $\quad (\dot{p}(t), p(t) - \int_t^1 \lambda(s)m(ds)) \in \partial g(\bar{x}(t), \dot{\bar{x}}(t)) + \{0\} \times N_K(\dot{\bar{x}}(t))$ a.e.

(10.6) $\quad p(1) \in \partial\alpha(x_1)$.

Remark: Certain closely related, but not strictly comparable results appear in the literature. The convex problem is treated by R.T. Rockafellar [18], who obtains necessary conditions couched in terms of vector measures and the Hamiltonian. Related cases are treated by J. Warga [20], H. Halkin [12], and F.H. Clarke [9, §6]. In these, no relation of the form (10.6) is obtained.

Proof. We may suppose that $x_o = 0$. Note that the solutions x of (10.2), (10.3) are bounded, so there is no loss of generality in assuming Ω compact, and in supposing that g satisfies the growth condition (1.3) of §1 with $r = 0$. We denote A_o^1 the set of absolutely continuous $x: [0,1] \to R^n$ such that $x(0) = 0$ and $\dot{x}$ belongs to L^1. If the norm of x is taken as $\int_o^1 |\dot{x}|dt$, then the dual of A_o^1 may be identified with L^∞, with the duality pairing

$$\langle x, v\rangle = \int_o^1 \dot{x}v dt .$$

We shall apply Corollary 4 of §7 with the following identifications:

$$x = (x,y)$$
$$U = L^1 \times L^1$$
$$W = A_o^1 \times L^1$$
$$f(x,y) = \int_o^1 g(x,y)dt$$
$$Z = A_o^1$$
$$Y = \{z \in Z: z(t) \in \Omega\}$$
$$C(x,y) = \int_o^t y(s)ds$$

$$\gamma(x,y) = x(1)$$
$$T = R^n$$
$$S = \{x_1\}$$
$$L(x,y) = \dot{x} - y$$
$$V = L^1$$
$$E(x,y) = \{0\} \text{ if } y \in K_1, \ \phi \text{ otherwise,}$$

where K_1 is the set of y in L^1 satisfying $y(t) \in K$ a.e. It is routine to check that $(\bar{x},\bar{y}) = (\bar{x},\dot{\bar{x}})$ then solves the abstract problem of §7, and to verify the hypotheses of the theorem. The controllability hypothesis follows mainly (much as in §9) from the observation that if (u,v,t,z) is sufficiently small in $L^1 \times L^1 \times R^n \times Z$, then the element (x,y) of W defined by

$$y(t) = \dot{x}_o(t) + u(t) + t - \int_o^1 (u+v)dt$$
$$\dot{x}(t) = y + v$$

satisfies

$$y \in K_1 + u$$
$$x(1) = x_1 + t$$
$$C(x,y) \in \Omega + z .$$

We calculate

$$V^* = L^\infty$$
$$L_o^* p = [-\dot{p},-p] \quad \text{(in the sense of distributions)}$$
$$V_o^* = \{p \in L^\infty: \dot{p} \in L^\infty\}$$

$$C^* v = (0, v)$$

$$\beta^* p = -p(1)$$

$$H(x,y,p) = 0 \text{ if } y \in K_1, = -\infty \text{ otherwise.}$$

A straightforward interpretation of (7.9) (with the help of Theorem 2, §6) then yields the existence of an arc p and elements τ, σ of $N_Y(C(\bar{x},\bar{y}))$ and $N_{K_1}(\bar{y})$ such that

$$(\dot{p}, p) \in \partial g(\bar{x}, \dot{\bar{x}}) + (0, \tau + \sigma)$$

$$p(1) \in \partial a(x_1)$$

Now, σ is a function in L^∞ satisfying

$$\int_0^1 \sigma(t) \cdot \{k(t) - \dot{\bar{x}}(t)\} dt \le 0$$

whenever k in L^1 is such that $k(t) \in K$ a.e. It follows readily from the measurable selection theorem [19, Theorem 4.1] that for almost every t, for every k in K,

$$\sigma(t) \cdot \{k - \dot{\bar{x}}(t)\} \le 0 ,$$

i.e. $\sigma(t) \in N_K(\dot{\bar{x}}(t))$. It remains now to characterize τ. We have by definition that for any arc x satisfying $x(0) = 0$, $x(t) \in \Omega$,

$$\int_0^1 \tau(t) \cdot \{\dot{x}(t) - \dot{\bar{x}}(t)\} dt \le 0 ,$$

where, as explained earlier, τ is identified with an element of L^∞. The lemma below then yields an expression for τ that concludes the proof.

Q.E.D.

We denote A^∞ the set of absolutely continuous functions $x: [0,1] \to R^n$ with derivative $\dot{x}$ in L^∞; p is an scalar ≥ 1.

Lemma 10.1 Let Ω be a compact convex set in R^n containing 0 in its interior, and let τ belong to $L^p([0,1]; R^n)$. If $\bar{x}$ in A^∞ solves the problem of maximizing

$$\int_0^1 \tau(s) \cdot \dot{x}(s) ds$$

over the elements x of A^∞ satisfying $x(0) = 0$, $x(t) \in \Omega$ for all t, then there is a function λ in L^∞ and a nonnegative Radon measure m on $[0,1]$ such that:

(10.7) $$\lambda(t) \in N_\Omega(\bar{x}(t)) \quad m\text{-a.e.}$$

(10.8) $$\tau(t) = \int_{[t,1]} \lambda(s) m(ds) \text{ a.e.}$$

Proof. Let h denote the support function of Ω; i.e.

$$h(p) = \max\{p \cdot v : v \in \Omega\} .$$

If S denotes the unit sphere of R^n, then x belongs to Ω iff $g(x) \le 0$, where

$$g(x) = \max\{p \cdot x - h(p) : p \in S\} .$$

The function g is locally Lipschitz, and it follows from [6, Theorem 2.1] that when $g(x) = 0$, $\partial g(x) \subset N_\Omega(x)$. Further, because $0 \in \text{int}\, \Omega$ we have $h(p) \ge \delta > 0$ when $p \in S$, and this implies $0 \notin \partial g(x)$ when $g(x) = 0$. The statement of the lemma implies then that $\bar{x}$ minimizes

14

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

MRC-TSR- REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
1900

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

9 Technical

4. TITLE (and Subtitle)

6 SHADOW PRICES, DUALITY AND GREEN'S FORMULA FOR A CLASS OF OPTIMAL CONTROL PROBLEMS.

5. TYPE OF REPORT & PERIOD COVERED
Summary Rep[redacted]t, no specific reporting period

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
10 J. P. Aubin [redacted]d F. H. Clarke

8. CONTRACT OR GRANT NUMBER(s)
15 DAAG29-75-C-0024
NRC-A 9082

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Mathematics Research Center, University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
#5-Mathematical Programming and Operations Research

11. CONTROLLING OFFICE NAME AND ADDRESS
See Item 18 below

12. REPORT DATE
11 Nov[redacted]r [redacted]978

13. NUMBER OF PAGES
43

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
12 26p.

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Research Council of Canada
Vancouver, B. C. Canada V6T 1W5

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
maximum principle, sensitivity, infinite horizon, generalized gradients

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A class of optimal control problems is considered in which the cost functional is locally Lipschitz (not necessarily convex or differentiable) and the dynamics linear and/or convex. By using generalized gradients and duality methods of functional analysis, necessary conditions are obtained in which the dual variables admit interpretation as shadow prices (or rates of change of the value function). Applications are presented in three settings: infinite-horizon optimal control, optimal control of partial differential equations and a variational problem with unilateral state constraints. A theorem is proved which characterizes the generalized gradients of integral functionals on L^p.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

221200

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

LB

$$\int_0^1 -\tau \cdot \dot{x}\, dt$$

subject to $x(0) = 0$, $g(x(t)) \le 0$. We may apply [9, Theorem 4] to deduce the existence of an arc p, a function λ in L^∞ and a Radon measure m supported on the set $\{t: g(\bar{x}(t)) = 0\}$ such that

$$(\dot{p}, p) + (0, \int_{(0,t)} \lambda dm) = (0, -\tau) .$$

Because $x(1)$ is free, the transversality condition

$$-p(1) = \int_{(0,1]} \lambda dm$$

also pertains. Thus p is identically equal to this last quantity, and the result follows.

Q.E.D.

REFERENCES

1. J.P. Aubin, Théorème du minimax pour une classe de fonctions, Comptes Rendus Acad. Sci. Paris 274 Série A (1972) 455-458.
2. __________, "Approximation of elliptic boundary-value problems", Wiley-Interscience, New York (1972).
3. J.P. Aubin, F.H. Clarke, Multiplicateurs de Lagrange en optimisation non convexe et applications, Comptes Rendus Acad. Sci. Paris 285 (1977) 451-454.
4. A. Bensoussan, E.G. Hurst Jr., B. Naslund, "Management Applications of modern control theory", North-Holland, American Elsevier, Amsterdam (1974)
5. N. Bourbaki, "Integration", Fascicule XXV, Chapitre VI, Hermann, Paris
6. F.H. Clarke, Generalized gradients and applications, Trans. Amer. Math. Soc. 205 (1975) 247-262.
7. __________, A new approach to Lagrange multipliers, Math. Oper. Research 1 (1976) 165-174.
8. __________, Multiple integrals of Lipschitz functions in the calculus of variations, Proc. Amer. Math. Soc. 64 (1977) 260-264.
9. __________, Generalized gradients of Lipschitz functionals, Tech. Report 1687, Math. Res. Center, Madison, Univ. of Wisconsin 1976 (Advances in Mathematics, to appear).
10. N. Christopeit, Necessary optimality conditions with application to a variational problem, SIAM J. Control and Optimization 15 (1977) 683-698.

11. H. Halkin, Necessary conditions for optimal control problems with infinite horizons, Econometrica 42 (1974) 267-272.

12. ________, Optimization without differentiability, in "Proceedings of the Conference on Optimal Control Theory" (Canberra, August 1977) Springer-Verlag, New York (to appear).

13. T.C. Koopmans, Concepts of optimality and their uses, Nobel Lectures 1975; reproduced in Amer. Econ. Rev. 67 (1977) 261-274.

14. G. Lebourg, Valeur moyenne pour gradient généralisé, Comptes Rendus Acad. Sci. Paris 281 Série A (1975) 795-797.

15. J.L. Lions, E. Magenes, "Problèmes aux limites non homogènes et applications", Dunod, Paris (1968).

16. R. Pallu de la Barrière, On the cost of constraints in dynamical optimization, in "Mathematical Theory of Control" (Ed. A.V. Balakrishnan, L.W. Neustadt), Academic Press, New York (1967).

17. R.T. Rockafellar, "Convex Analysis", Princeton Univ. Press, Princeton, N.J. (1970).

18. ______________, State constraints in convex control problems of Bolza, SIAM. J. Control and Optimization 10 (1972) 691-715.

19. D.H. Wagner, Survey of measurable selection theorems, SIAM J. Control and Optimization 15 (1977) 859-903.

20. J. Warga, Controllability and necessary conditions in unilateral problems without differentiability assumptions, SIAM J. Control and Optimization 14 (1976) 546-572.